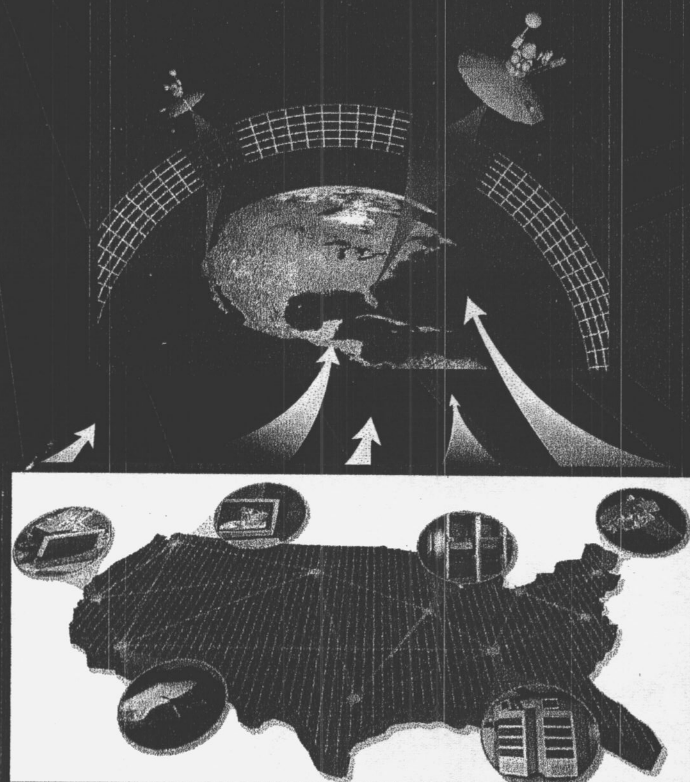
arc- IPG_ 30x24.ai

Source of Acquisition
NASA Ames Research Center

NASA
Ames Research Center
www.arc.nasa.gov

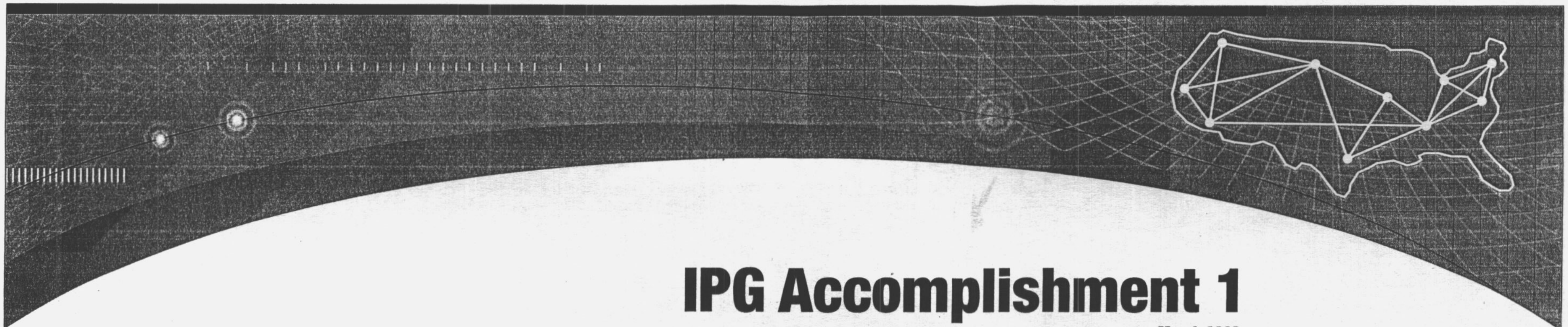
Information Power Grid (IPG)



Grids are an emerging technology that provide seamless and uniform access to the geographically dispersed computational, data storage, networking, instruments, and software resources needed for solving large-scale scientific and engineering problems.

NASA's Information Power Grid (IPG) project is developing and deploying such a computing and data grid. Its goal is to use NASA's remotely located computing and data system resources to build distributed systems that can address problems too large or complex for a single site.

The IPG provides a persistent infrastructure and will, on demand, locate and co-schedule the resources needed to solve large-scale or widely distributed problems. IPG services will support workflow management frameworks that coordinate the processes inherent in distributed science and engineering problems of the NASA Enterprises: simulation, design, data collection, monitoring and control. IPG is a multicenter effort in collaboration with universities and other goverment agencies.



IPG Accomplishment 1

March 2000

Access to Distributed Data

DESCRIPTION:
A key function of Grids is to provide uniform access to widely distributed resources, including heterogeneous distributed archival data and information systems. On-demand access to widely distributed archived data and information enhances engineering and scientific collaboration.

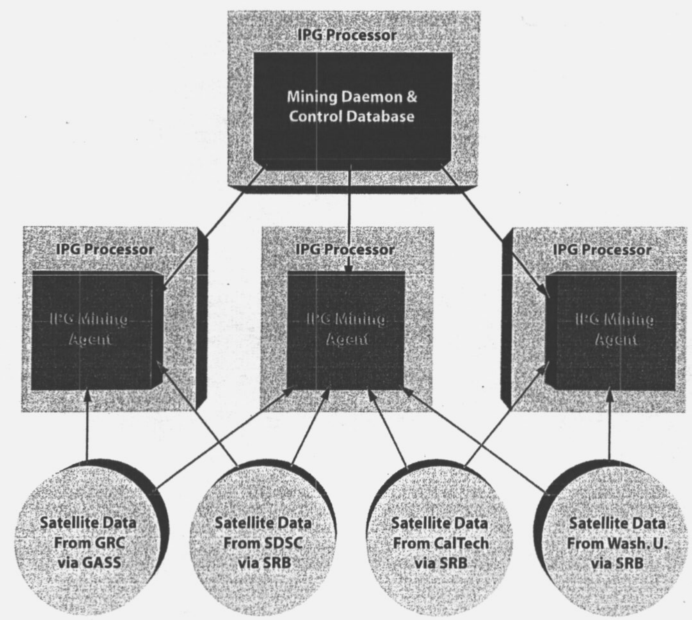
The objective of this accomplishment was to develop software that enables seamless access to catalogued archival data and information distributed throughout multiple NASA Centers and IPG collaborator sites.

Access to distributed data, a capability built into the IPG infrastructure, integrates archival storage systems at NASA Ames Research and NASA Goddard Space Flight Centers, and the Jet Propulsion Laboratory. IPG/Globus security mechanisms are also integrated into the IPG environment to provide strong access control and secure remote data access.

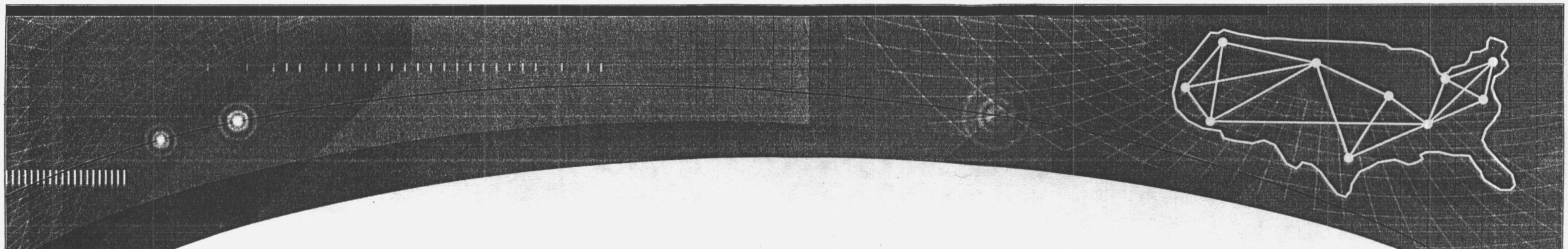
DEMONSTRATION:
The Grid Miner, a data mining application, runs on multiple data archive sites within the IPG environment. Data are accessed using a metadata catalogue and uniform access methods — Storage Resource Broker (SRB) and Globus Access to Secondary Storage (GASS), which achieve high-speed remote data access. To extract knowledge from remotely sensed satellite data, the Grid Miner uses computationally intensive techniques — subsetting and analysis of ocean surface temperature data, for example.

The Grid Miner uses agents and the IPG to:

- Stage mining agent(s) and a mining plan to one or more IPG processors
- Consult the Mining Information Server to identify the Mining Operator Repository appropriate for the type of IPG processor to be used for mining
- Download necessary mining operators from the Operator Repository to support a mining plan
- Acquire from the Mining Database the repository addresses of data to be mined
- Download data to be mined from IPG-accessible data repositories
- Mine data by applying the mining operators specified in the mining plan
- Produce one or more output files with the mining results



www.ipg.nasa.gov



IPG Accomplishment 2

September 2000

IPG Heterogeneous Computing

DESCRIPTION:
The IPG provides middleware services for building large-scale, dynamically constructed problem solving environments from distributed, heterogeneous resources. The objective of this accomplishment was to develop a prototype heterogeneous, distributed computing environment.

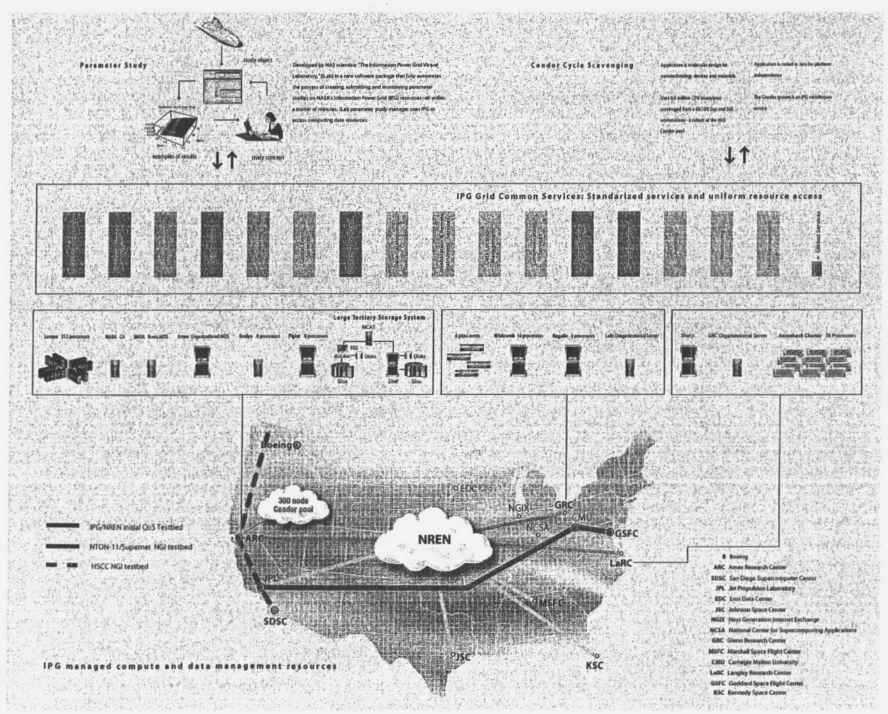
System tools and software provided include linked testbeds integrating two or more different types of machine architectures at three NASA centers. The prototype computing environment was tested using parameter study and cycle scavenging applications.

IPG software services for resource discovery, uniform access to geographically and organizationally dispersed computing and data resources, job management, single sign-on, security, inter-process communication, and resource management were also provided in the linked testbeds.

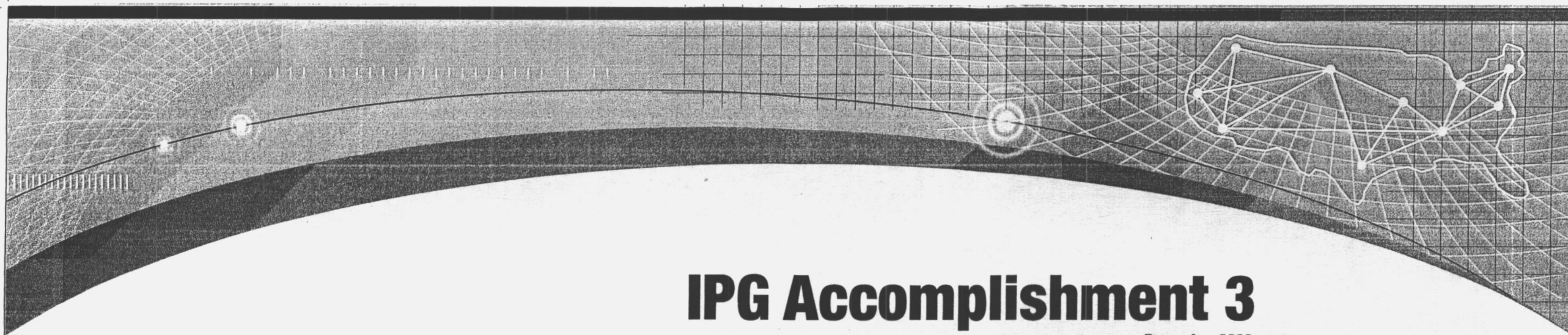
DEMONSTRATION:
Both the services and the operational support associated with the linked testbeds are in place at NASA Ames, Glenn, and Langley Research Centers.

Hardware resources for the baseline IPG prototype-production system include approximately 600 CPU nodes in six SGI Origin 2000s distributed across the three NASA centers, 100 Terabytes of uniformly and securely accessible mass storage, several workstation clusters with about 100 CPUs collectively, and a Condor pool of 200 workstations.

The outcome of this accomplishment is a reduction in end-to-end turnaround time for aerospace simulation problems, specifically peak performance and cost performance. The parameter study system, Ilab, used to complete this demonstration provides a highly efficient method for studying complex systems, resulting in reduction of turnaround time for system simulation. The molecular design application coded in Java and managed by the Condor cycle scavenger applies hundreds of thousands of megaflop years of otherwise idle computing time to solving significant NASA problems.



www.ipg.nasa.gov



IPG Accomplishment 3

December 2000

Integration of Large-Scale Computing Node into Distributed Environment

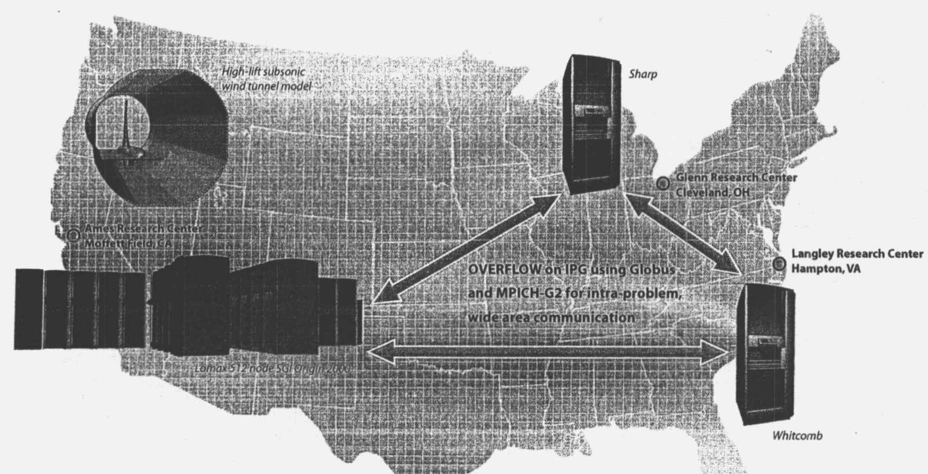
DESCRIPTION:
The goal of this accomplishment was to acquire and incorporate new large-scale computing systems into NASA's heterogeneous distributed computing environment, IPG, and demonstrate its ability to execute seamless operations. Incorporating a large-scale computing node into the IPG environment provides the capability to generate solution datasets for a number of large and complex aerospace research problems. This results in a reduction in end-to-end turnaround time for aerospace design and simulation problems.

IPG's distributed computing environment provides the middleware services for building large-scale, dynamically constructed problem solving environments from distributed, heterogeneous resources.

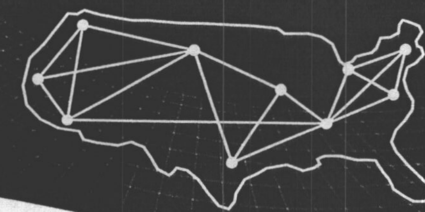
DEMONSTRATION:
IPG's basic Grid services and their operational support are in place at NASA Ames, Glenn, and Langley Research Centers. In addition to other computing and storage resources distributed across the three NASA centers (uniformly and securely accessible mass storage, several PC clusters with about 100 CPUs collectively, and a Condor pool of 200 workstations), the IPG prototype-production system now includes a 512 node SGI Origin 2000, Lomax.

Application POC: Mohammad J. Djomehri

The accompanying illustration demonstrates several potential improvements in the aerospace design milieu, all of which contribute to reducing design time. First, IPG provides a uniform view of resources, including a large-scale computing node, providing computational scientists with a broader range of resources to solve problems — the same interface can be used to access all of the resources noted above. Second, large-scale distributed systems can be built by using the IPG mechanisms for locating and aggregating components, potentially providing the ability to solve larger problems not previously possible.



The research branch of NAS is investigating algorithms that are suitable for a Grid computing "meta-platform." One candidate is overset grid codes that can tolerate timestep mis-matches on the intra-object boundaries. A version of the OVERFLOW, Navier-Stokes, CFD simulation code is being modified to test this approach. It has been demonstrated operating across systems at ARC, GRC, and LaRC, solving for flow about large test objects mounted in a wind tunnel.



IPG Accomplishment 4

September 2001

Remote Access to High Data-Rate Instruments

DESCRIPTION:

IPG's distributed computing environment provides the middleware services for building large-scale, dynamically constructed problem solving environments using distributed, heterogeneous resources. The IPG also provides a uniform security infrastructure for users to access computing and data storage resources to run applications. IPG's basic Grid services and accompanying operational support are in place at NASA Ames, Glenn, and Langley Research Centers and the Jet Propulsion Laboratory.

The goal of this accomplishment was to demonstrate the use of Grid services for remote connectivity to high data-rate instruments and distributed real-time access to instrument data. Specifically, to provide connectivity to one or more remote instrument from the IPG environment at a data access rate exceeding 50 Mbits/sec, demonstrating the use of Grid services and infrastructure for on-demand connectivity to high data-rate instruments and enhanced engineering and scientific collaboration.

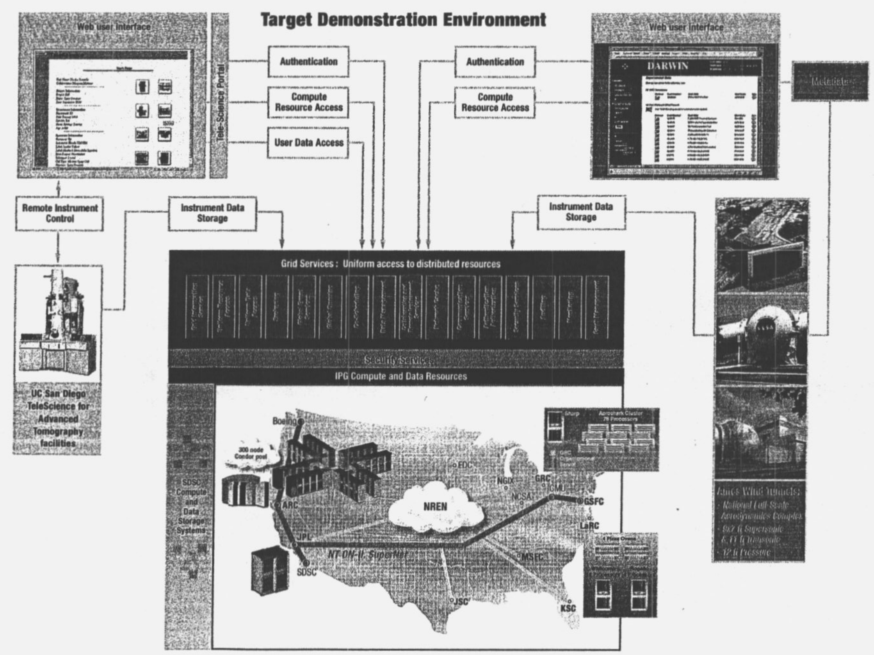
DEMONSTRATION:

Two scenarios were used to demonstrate remote access to high data-rate instruments:

1. The Ames DARWIN/DREAM wind tunnel instrumentation system used IPG services and resources in order to: a) provide the instrumentation facility access to large-scale computing and data systems; b) provide a standardized set of highly capable resource access and management services that do not have to be developed by the application developers, and; c) provide wider access to instrumentation data and analysis systems through a widely deployed IPG.

2. UC San Diego's (UCSD) TeleScience for Advanced Tomography facilities were used to demonstrate that NASA researchers using remote, non-NASA facilities, have the ability to use NASA IPG services and resources to both store and analyze data obtained from remote instrumentation facilities.

The accompanying illustration represents the two instrument systems used — one at Ames and one at UCSD, both of which use Grid services to provide secure data management and remote access to various capabilities of the instruments. DARWIN system users were scattered across the country, while the TeleScience system had a NASA user at Wallops Island manipulating the instrument at UCSD. Both systems stored data on IPG resources at Ames. All of the critical data paths for both demonstrations transferred data at 50 Mbits/sec, or greater.



IPG POC: William E. Johnston, www.ipg.nasa.gov

IPG Accomplishment 5

September 2002

Exploratory Grid Environment

DESCRIPTION:
One of the IPG team's goals during 2002 was to develop an exploratory Grid (IPG) environment to provide location-independent, heterogeneous compute and data storage resources, and to support access to distributed datasets and high-confidence design and simulation tools for NASA scientists and engineers. This prototype IPG environment will be used for a new problem-solving paradigm for NASA.

DEMONSTRATION:
An IPG prototype environment was developed to provide consistent access to heterogeneous resources using software installed on all IPG systems, and operated using common procedures. To test this environment, IPG-based job launcher and run manager software packages were developed to automatically launch, execute, monitor, restart, terminate, enter into a database, and analyze a large number of simulations over a short period of time.

The IPG team worked on development of 17 computer and data storage resources distributed over seven locations. These heterogeneous compute systems run Irix, Solaris, and Linux operating systems.

NASA Advanced Supercomputing (NAS) Division scientists utilized 14 of these resources at four locations for simulations of a selected NASA CICT/CNIS (Computing, Information, and Communication Program under the Computing, Networking, and Information Systems Project) Grand Challenge Application on the IPG exploratory Grid.

Used to analyze a flight vehicle over a large range of flight conditions, the AeroDB application provides an automated system for executing Computational Fluid Dynamics (CFD) programs. AeroDB was used to implement and monitor a parameter study of a Liquid Glide-Back Booster (LGBB) vehicle consisting of 2,863 Cart3D and 211 Overflow runs executed in seven consecutive days. The demonstration confirmed a new problem solving paradigm capability.

www.ipg.nasa.gov

CICT/CNIS Program Grand Challenge Applications

NASA IPG Nodes:
NAS Division, Ames Research Center
AS Division, Ames Research Center
Glenn Research Center
Langley Research Center

Collaborating Nodes:
National Center for Supercomputing Applications (NCSA)
University of Southern California - ISI
Argonne National Laboratory